

Orbiting Carbon Observatory-2 (OCO-2) cloud screening; validation against collocated MODIS and initial comparison to CALIOP data

Thomas E. Taylor¹, Christopher W. O'Dell¹, Christian Frankenberg², Philip Partain³, Heather W. Cronk³, Andrey Savtchenko⁴, Robert R. Nelson¹, Emily J. Rosenthal¹, Albert Chang², David Crisp², Annmarie Eldering², and Mike Gunson²

¹Department of Atmospheric Science, Colorado State Univ., Fort Collins, CO, USA.

²JPL, Pasadena, CA, USA.

³Cloudsat Data Processing Center, Cooperative Institute for Research in the Atmosphere, Fort Collins, CO, USA.

⁴ADNET, NASA Goddard Space Flight Center, Code 610.2, Greenbelt, Maryland

Correspondence to: T.E. Taylor (tommy.taylor@colostate.edu)

Abstract. The retrieval of the column-averaged carbon dioxide (CO₂) dry air mole fraction (X_{CO_2}) from satellite measurements of reflected sunlight in the near-infrared can be biased due to contamination by clouds and aerosols within the instrument's field of view (FOV). Therefore, accurate aerosol and cloud screening of soundings is required prior to their use in the computationally expensive X_{CO_2} retrieval algorithm. Robust cloud screening methods have been an important focus of the retrieval algorithm team for the National Aeronautics and Space Administration (NASA) Orbiting Carbon Observatory-2 (OCO-2), which was successfully launched into orbit on July 2, 2014.

Two distinct spectrally-based algorithms have been developed for the purpose of cloud clearing OCO-2 soundings. The A-Band Preprocessor (ABP) performs a retrieval of surface pressure using measurements in the 0.76 μm O₂ A-band to distinguish changes in the expected photon path length. The Iterative Maximum A-Posteriori (IMAP) Differential Optical Absorption Spectroscopy (DOAS) (IDP) algorithm is a non-scattering routine that operates on the O₂ A-band as well as two CO₂ absorption bands at 1.6 μm (weak CO₂ band) and 2.0 μm (strong CO₂ band) to provide band-dependent estimates of CO₂ and H₂O. Spectral ratios of retrieved CO₂ and H₂O identify measurements contaminated with cloud and scattering aerosols.

Information from the two preprocessors is feed into a sounding selection tool to strategically down select from the order one million daily soundings collected by OCO-2 to a manageable number (order 10 to 20%) to be processed by the OCO-2 L2 X_{CO_2} retrieval algorithm. Regional biases or errors in the selection of clear-sky soundings will introduce errors in the final retrieved X_{CO_2} values,

ultimately yielding errors in the flux inversion models used to determine global sources and sinks of CO_2 .

In this work collocated measurements from NASA’s Moderate Resolution Imaging Spectrometer (MODIS), aboard the *Aqua* platform, and the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP), aboard the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) satellite, are used as a reference to access the accuracy and strengths and weaknesses of the OCO-2 screening algorithms. The combination of the ABP and IDP algorithms is shown to provide very robust and complimentary cloud filtering as compared to the results from MODIS and CALIOP. With idealized algorithm tuning to allow throughputs of 20-25%, correct classification of scenes, i.e., accuracies, are found to be $\simeq 80\text{-}90\%$ over several orbit repeat cycles in both the winter and spring time for the three main viewing configurations of OCO-2; nadir-land, glint-land and glint-water.

Investigation unveiled no major spatial or temporal dependencies, although slight differences in the seasonal data sets do exist and classification tends to be more problematic with increasing solar zenith angle and when surfaces are covered in snow and ice. An in depth analysis on both a simulated data set and real OCO-2 measurements against CALIOP highlight the strength of the ABP in identifying high, thin clouds while it often misses clouds near the surface even when the optical thickness is greater than 1. Fortunately, by combining the ABP with the IDP, the number of thick low clouds passing the preprocessors is partially mitigated.

1 Introduction

NASA’s OCO-2 satellite was launched on July 2, 2014 into a sun-synchronous orbit. After an initial on-orbit satellite bus checkout period it was inserted into the Afternoon Constellation, i.e., the A-Train (Stephens et al., 2002), from where it will collect measurements of reflected solar radiation in tandem with the other A-Train sensors such as MODIS, CloudSat and CALIOP. The OCO-2 instrument contains three bore sighted telescopes, the light from which is dispersed via gratings to form two dimensional images of spectra onto a 1024×1024 pixel focal plane array (Crisp et al., 2004). The three spectral bands, centered at $0.76 \mu\text{m}$ (O_2 A-band), $1.61 \mu\text{m}$ (weak CO_2 band) and $2.0 \mu\text{m}$ (strong CO_2 band) were chosen to provide the greatest constraint possible on obtaining high precision retrievals of X_{CO_2} . From an altitude of approximately 700 km, and with an integration time of 3 Hz, each OCO-2 footprint is approximately 2 km cross-track and 3 km along-track. The orientation of the satellite bus rotates with latitude to align the optical elements at a constant orientation relative to the principle scattering plane defined by the earth-sun-satellite geometry.

For heavily contaminated scenes, the full physics X_{CO_2} retrieval algorithm will likely fail, i.e., non-convergence of the optimizer, thus wasting valuable processing time. Worse, contamination by even low optical thicknesses ($\lesssim 0.3$) can introduce scene-dependent biases in the X_{CO_2} , hindering

the ability to accurately determine the sources and sinks on regional scales - the primary mission of OCO-2.

It is therefore necessary to provide accurate cloud screening on all of the approximately one million OCO-2 measurements collected on a daily basis. The OCO-2 sampling approach was designed to mitigate the chances of introducing systematic biases in the retrieved X_{CO_2} values (Crisp et al., 2004). Two primary mitigation strategies related to cloud screening are the satellite's multiple observation modes and the small native footprint size of the instrument's field of view (FOV). As discussed in Miller et al. (2007), nadir viewing observations, with the instrument bore sighted directly beneath the satellite orbit track, minimizes the FOV size of individual footprints but also yields low signal to noise ratios (SNR) over water surfaces, which are very dark in the CO_2 channels. Observations in glint viewing mode, with the bore sight oriented towards the point of specular reflection, maximizes the SNR, but yields larger footprint sizes, increasing the likelihood of cloud contamination within the FOV. The operational viewing strategy of OCO-2 in the early phase of the mission (Sep-2014 through Jun-2015) was to alternate subsequent sixteen day repeat cycles in nadir and glint viewing geometries. However, on 2-Jul-2014, nominal sequence was modified to alternate each orbit between nadir and glint, with plans to designate certain orbit paths containing a very large fraction of ocean as always glint.

Note that OCO-2 also operates in the so called target mode, which constantly reorients the spacecraft to view a small area as it approaches, passes overhead and recedes from a ground target site. Target sites are typically home to one of the TCCON stations, which allows for X_{CO_2} validation and air mass dependency bias studies (Wunch et al., 2010, 2011). Nominally one orbit each day is designated as a target orbit, with acquisition made only if the skies are determined to be relatively clear and while the local target solar zenith angle (SZA) is less than approximately 55° . Early analysis of target orbits is beyond the scope of this work, but discussion can be found in (Wunch et al., 2015).

Prior to the launch of OCO-2, the algorithm development team had the benefit of working with the Japanese Aerospace Exploration Agencies (JAXA) Greenhouse gas Observation Satellite (GOSAT) data set (Kuze et al., 2009; Yoshida et al., 2011). Analysis of the ABP cloud screening performance, similar to that presented here, was published in (Taylor et al., 2012), who concluded that cloud screening accuracies relative to MODIS were around 80% over land and 90% over ocean surfaces for the ABP. Note that the IDP algorithm was not available at that time. The GOSAT had the added benefit of the on-board Cloud and Aerosol Imager (CAI), which also compared favorably against MODIS (Ishida et al., 2011) when both data sets were processed through the Cloud and Aerosol Unbiased Decision Intellectual Algorithm (CLAUDIA) developed as the primary cloud screening algorithm by Japan's National Institute of Environmental Studies (NIES) (Ishida and Nakajima, 2009).

The work presented here shows for the first time a validation of the OCO-2 cloud screening applied to a set of real measurements collected during the first year of operations after completion of the on-

orbit instrument checkout (IOC) phase. The data set for analysis is comprised of four full repeat cycles, two in nadir and two in glint viewing, from both a winter time and spring time range. This provides a statistically robust global analysis of the OCO-2 cloud screening performance.

95 The organization of the work presented in this manuscript is as follows; In Section 2, the two OCO-2 cloud screening algorithms are discussed and detailed analysis of performance on simulated data is provided. Section 3 introduces the collocated MODIS and CALIOP products, which will provide a valuable resource for many scientific investigations and will be made available to researchers via the Jet Propulsion Laboratory (JPL) CO₂ data portal. Section 4 provides detailed analysis of
100 the cloud screening validation procedure, including optimization of algorithm tuning and the direct comparison against both MODIS and CALIOP. Finally, summary conclusions and ideas for future lines of investigation are given in Section 5.

2 OCO-2 Aerosol and Cloud Screening Algorithms

The OCO-2 A-Band Preprocessor (ABP) and the Iterative Maximum A-Posteriori (IMAP) Differential Optical Absorption Spectroscopy (DOAS) (IDP) algorithms are applied to the full OCO-2 data
105 set by the Science Data Operations System (SDOS) at JPL (**Is there an actual citation for SDOS?**). Since this is on the order of one million soundings per day, both algorithms are made computationally efficient by neglecting atmospheric scattering in the radiative transfer forward model (although ABP does account for Rayleigh scattering by oxygen molecules which is non-negligible in the O₂
110 A-band). By assuming clear-sky conditions, deviations of retrieved variables from expected values allows for the identification of scenes contaminated by cloud and aerosol. Brief explanations of both algorithms are given below. In addition we provide a detailed discussion of the merits of combining the two into a single cloud and aerosol filter and directly compare the performance on a set of simulated OCO-2 L1B radiances.

115 2.1 The A-Band Preprocessor (ABP)

The ABP algorithm was introduced and applied to early GOSAT data in Taylor et al. (2012), with further analysis performed on realistic GOSAT simulations given in O'Dell et al. (2012). More detail and optimization testing can be found in the Algorithm Theoretical Basis Document available on the NASA OCO-2 web home page (O'Dell et al., 2014).

120 The ABP algorithm employs a Bayesian retrieval to estimate surface pressure and surface albedo from high resolution spectra in the 0.765 μm O₂ A-band. The absorption by oxygen molecules in this spectral region (approximately 0.759 to 0.771 μm) produces a signature in the measured reflectances that can be inverted via standard retrieval techniques to provide estimates of atmospheric parameters contained in the state vector. Using some prior knowledge of the expected values, these parameters
125 can be interpreted to provide information on contamination within the FOV of the satellite sensor.

The radiative transfer forward model (FM) assumes a clear-sky condition, i.e. Rayleigh scattering only, such that differences between the modeled and measured radiances are apparent when the scene contains cloud or aerosol. Estimates of the surface pressure, differenced against the apriori ECMWF reanalysis value; $\Delta p_{s,cld} = p_s - p_{s,a}$, surface albedo (α) and the χ^2 goodness-of-fit statistic are used in conjunction with the instrument signal-to-noise ratio (SNR) to reveal changes in the expected optical path length, allowing scenes to be flagged as cloudy or clear.

Simulations have demonstrated the ability of the ABP to accurately determine scenes contaminated with mid- or high- altitude clouds, although it has trouble detecting low level clouds even when they are optically thick (O'Dell et al., 2012).

2.2 The IMAP-DOAS Preprocessor (IDP)

The IDP algorithm performs computationally efficient, non-scattering retrievals separately for the weak and strong CO₂ bands and then computes ratios of the retrieved CO₂ (r_{CO_2}) and H₂O (r_{H_2O}). Path length differences induced by clouds or aerosols yield deviations in the ratios from expected values near 1.0 (Frankenberg, 2014). Details of the mechanics of the algorithm can be found in (Frankenberg et al., 2005). In contrast to the ABP, which is more sensitive to the altitude of the effective scattering layer, the IDP is more sensitive to spectral variation associated with small scattering particles, especially aerosols.

Within IDP, each single-band retrieval determines column gas amounts by fitting simultaneously to scale factors applied to prior CO₂ and H₂O profiles. Additionally, a dispersion shift is fit for each band. Because the single band retrievals allows for independent constraint of each gas, the ratio of the gas retrievals from the two bands provides information on the presence of clouds or aerosols. This occurs primarily because of different surface albedo and particle scattering characteristics between the two bands. For most soundings, only reasonably clear scenes will yield ratios close to unity. Also, because CO₂ and H₂O have rather different scale heights, they are affected by a given cloud or aerosol layer differently and hence can provide somewhat independent information.

2.3 Combined ABP and IDP Filter on Simulated Data

In this section we test the effectiveness of the combined ABP and IDP filters using simulated data, similar to the work reported in Section 3.2 of O'Dell et al. (2012). That analysis showed that 20-40% of thick, low water clouds or aerosol layers with total optical depth (TOD) $\gtrsim 1$ can be missed by the O₂ A-band filter for GOSAT over land. The culprit appears to be a nearly-complete cancellation of photon path length shortening and lengthening, which can occur for certain combinations of cloud top pressure, cloud optical depth, solar zenith angle and the O₂ A-band surface albedo (e.g., see Section 2 of Taylor and O'Brien (2009)). This may be related to the "critical albedo" phenomenon recently described in Seidel and Popp (2012). In general, these cancellation effects can also occur in the weak and strong CO₂ bands, but are unlikely to occur in all three simultaneously.

Here we expand on that work using a new set of simulations for both GOSAT and OCO-2 calculated via the CSU orbit simulator model (O’Brien et al., 2009), which has realistic distributions of clouds, aerosols, surface types, and viewing geometries. The simulation data set consists of 96 orbits spanning 3 days in both June and December to cover a full range of solar zenith angles. Soundings with a sub-satellite point over land were set to nadir viewing geometry, while those over water were set to view the spectral glint spot. A temporal sampling of 1 Hz was used.

In essence the two data sets are identical with the exception of the instrument model used to convolve the TOA reflected radiances. The major differences in the GOSAT and OCO-2 instruments are polarization sensitivity, Instrument Line Shape (ILS) and the noise models. Full details on the specific sensors and calibration procedures can be found in Crisp and Johnson (2005), O’Dell et al. (2011), Day et al. (2011), Rosenberg et al. (2015), Lee (2015) and Boesch et al. (2015) for OCO-2 and Kuze et al. (2009) and Yoshida et al. (2010) for GOSAT.

Panel (a) of Figure 1 shows $\Delta p_{s,cld}$ versus the true cloud plus aerosol optical depth (total optical depth, TOD) for ~ 30 k synthetic soundings in nadir viewing mode over land in the month of June. The soundings are colored by the cloud relative height, defined as the height at which the partial-column TOD at 760 nm reaches the smaller of 50% of the TOD or unity, where the integration begins at the top of the atmosphere. Values near zero (dark colors) represent high cloud or aerosol layers, while values near unity (bright colors) represent low cloud or aerosol layers. The horizontal black lines show thresholds used to separate cloudy scenes from clear-sky. Each point on the plot is a 2-dimensional binned mean value to avoid over plotting of data.

For both OCO-2 and GOSAT, the value of $\Delta p_{s,cld}$ (upper panels) diverges from zero at a lower TOD for the high clouds (dark colors) than it does for low clouds (bright colors). This is an indicator of the ABP’s ability to detect high, optically thin clouds due to strong photon path length modification. The ABP’s struggle to identify low clouds - even those that are optically thick, is seen by the large number of bright colored data points with small $\Delta p_{s,cld}$ at high TOD.

For OCO-2 the divergence of $\Delta p_{s,cld}$ from zero for high cloud occurs at lower values of TOD than it does for GOSAT, indicating a higher sensitivity in the O₂ A-band. This implies that the ABP is better able to pick out lightly cloud contaminated scenes for OCO-2, while conversely, the X_{CO_2} L2 retrieval algorithm may be more sensitive to contamination by thin scattering layers for OCO-2 than it was for GOSAT. This is likely due to the OCO-2 noise model, which provides higher SNR in the absorption line cores relative to the continuum than does the spectrally flat noise model of GOSAT. The implication is that the threshold value of $\Delta p_{s,cld}$ can be set more loosely for OCO-2 than for GOSAT to achieve the same sounding pass rate.

The IDP CO₂ and H₂O ratios (designated r_{CO_2} and r_{H_2O} , respectively) are both correlated with the retrieved effective albedo ratios, as shown in panels (b) and (c) of Figure 1. For low TOD ($\lesssim 0.1$), the ratios are close to unity. However, for larger TOD cases, r_{CO_2} and r_{H_2O} tend to diverge from unity, almost exclusively in the positive direction. This is a result of the lower surface albedo over

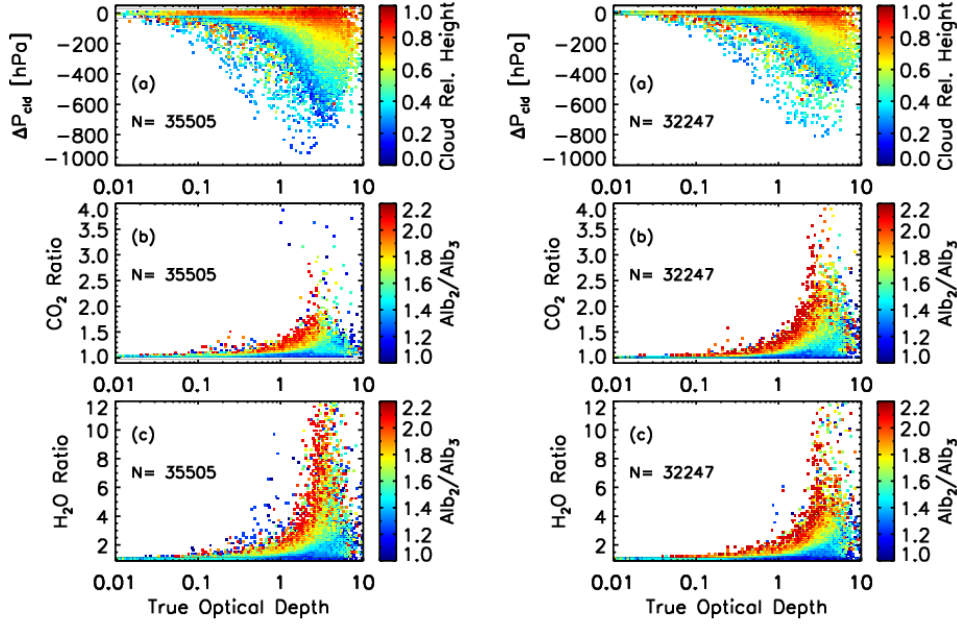


Figure 1. Scatter plots of (a) $\Delta p_{s,cld}$ (b) CO_2 ratio ($r\text{CO}_2$) and (c) H_2O ratio ($r\text{H}_2\text{O}$) versus true optical depth for OCO-2 (left) and GOSAT (right) simulations over land for the month of June. In panel (a), each sounding is colored according to the cloud relative height (see text), while in panels (b) and (c), each sounding is colored according to the ratio of the $1.6\ \mu\text{m}$ to the $2.0\ \mu\text{m}$ retrieved effective albedo. The horizontal black lines show the selected threshold values presented in Table 1.

land in the strong CO_2 band relative to the weak CO_2 band. This implies that path-shortening dominates in the strong CO_2 band, as photons forward-scattered through a cloud layer will be absorbed by the surface more often than in the weak CO_2 band. Thus, IDP retrievals in the strong CO_2 band will yield a smaller gas column (of either CO_2 or H_2O) than will those in the weak CO_2 band, yielding weak/strong ratios greater than unity.

As the albedo ratios diverge from 1.0 (move from dark to bright colors in the plots), the CO_2 and H_2O ratios also diverge from 1.0 for lower values of the true AOD, thus allowing for more effective screening. In other words, when there is significant spectral contrast in the scene, the IDP has higher fidelity in identifying contamination by cloud and aerosol.

A direct comparison between the OCO-2 to GOSAT IDP ratios reveals that there is little difference in $r\text{H}_2\text{O}$ for the two instruments (panels (c) in Figure 1), while for $r\text{CO}_2$ there is stronger divergence from 1 at lower TOD for GOSAT than for OCO-2 (panels (b) in Figure 1). This implies that the $r\text{CO}_2$ is a less sensitive indicator of cloud for OCO-2, consistent with recent findings (Lukas Mandrake; private communication).

The cloud screening results for the June land-viewing simulations are presented in Figure 2, which shows the fraction of soundings identified as clear by ABP only (blue), IDP only (green) and the

Table 1. Settings of the ABP and IDP filter thresholds used for the simulated data set.

	OCO-2	GOSAT
$\Delta p_{s,cld}$ (hPa)	30.0	10.0
χ^2 scale factor	5.0	2.0
r CO ₂ center point	1.0	1.0
r CO ₂ half-width	0.022	0.015
r H ₂ O center point	1.02	1.02
r H ₂ O half-width	0.10	0.10

combined set (black). The total number of scenes and the percent identified as clear are labeled on
215 each panel for the three filter combinations in the corresponding colors. Shown are all-scenes cases
in panels (a), high cloud cases in panels (b) and low clouds in panels (c). The filters were tuned
using the all-scenes case to produce approximately the same pass rate for both OCO-2 (left column)
and GOSAT (right column). The values of the selected screening variables for both instruments are
provided in Table 1. Note the more aggressive $\Delta p_{s,cld}$, χ^2 multiplier and r CO₂ threshold settings
220 for GOSAT relative to OCO-2 in order to achieve the same sounding throughput.

The top panels indicate general agreement between the two sensors in the all-scenes case. For ex-
ample, when TOD = 0.25, 50% and 61% of the scenes are being identified as clear by the combined
filters for OCO-2 and GOSAT, respectively. It is worth noting that the combination of ABP and IDP
provides a more aggressive screening than a single filter, indicating that they are keying off different
225 features in the spectra and not flagging identical soundings. For both OCO-2 and GOSAT all three
filter combinations exhibit a smooth decay toward 0.0 fraction passing with increasing TOD, except
for a noticeable “hump” in the fraction identified as clear when TOD \simeq 3, i.e., a misidentification of
cloudy scenes as clear. This anomaly is much more pronounced for OCO-2 than GOSAT.

To investigate these missclassifications in more detail, the same analysis was performed on two
230 sounding subsets; high-clouds shown in panels (b) and low clouds in panels (c), where high (low)
is defined as cases where 95% of the TOD resides in the top 40% (bottom 30%) of the atmosphere.
About 4% and 18% of the total number of soundings were found to contain high cloud and low
cloud, respectively.

For the GOSAT high clouds case, we see that at TOD = 0.25 the percent identified as clear is 3%,
235 47% and 3% for the ABP, IDP and combined filters, respectively. This is consistent with previous
results (Taylor et al., 2012; O’Dell et al., 2012) which have shown the ABP filter to be extremely
effective at screening high, thin clouds. In comparison, for OCO-2 the high clouds case the corre-
sponding values are 0%, 4% and 0%.

The lower panels of Figure 2 show that a much greater fraction of the low clouds are passing
240 for OCO-2 than for GOSAT. For example, when TOD = 1.0, 83%, 64% and 61% of scenes pass

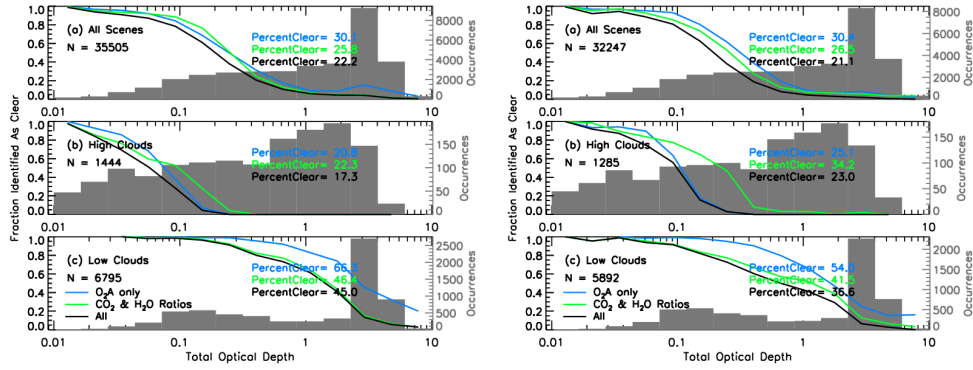


Figure 2. The fraction of simulated OCO-2 (left) and GOSAT (right) soundings identified as clear by the ABP screen alone (blue), by the $r\text{CO}_2$ plus $r\text{H}_2\text{O}$ (green), and by all three filters combined (black), plotted as a function of the total cloud plus aerosol optical depth (TOD) at 760 nm. The frequency histogram of the TOD is plotted in gray against the right ordinate. Panel (a) shows all cloud cases, while panel (b) shows only those scenes where 95% of the OD resides in the upper 40% of the atmosphere (i.e., high clouds) and panel (c) shows cases where 95% of the OD resides in the lowest 30% of the atmosphere (i.e., low clouds layers). Only the June nadir-land data is shown.

the ABP, IDP and combined filters, respectively, for OCO-2, while only 66%, 52% and 41% pass for GOSAT. This supports the findings presented in (O'Dell et al., 2012) for GOSAT and confirms that ABP alone is unlikely to detect low clouds observed by OCO-2, but that by combining the two filters, the number of low cloud contaminated scenes is reduced by about 50% over using the ABP alone.

In summary, combining the ABP and IDP cloud filters in tandem, yields a robust cloud and aerosol screener that is more effective at identifying scenes with both high- and low-altitude scattering material than either algorithm alone. A statistical χ^2 value from the ABP fit, the ABP $\Delta p_{s, \text{cld}}$ and the IDP $r\text{CO}_2$ and $r\text{H}_2\text{O}$ values serve as some of the primary inputs into the OCO-2 scene selection algorithm (Mandrake et al., 2013), which will be used to down select to approximately 10-20% of the total measured soundings to be processed by the L2 X_{CO_2} retrieval algorithm. Furthermore, the Data Ordering Genetic Optimization (DOGO) algorithm will be used to assign to each converged sounding a “warn level” value (an integer between 1-20) to quantify its likelihood of providing an unbiased estimate of X_{CO_2} (Mandrake et al., 2015).

3 OCO-2, MODIS and CALIOP collocated data sets

NASA’s Earth Observing System (EOS) Afternoon Constellation (the A-Train) is comprised of a suite of satellites flying in tight formation that provide near-simultaneous observation from a multitude of sensors (Stephens et al., 2002). The wide swath (2330 km cross-track) of the MODerate

Table 2. Summary of OCO-2 B7 data set to which MODIS and CALIOP collocation were performed.

Index	View Mode	Start Date	End Date	NumDays	NumOrbits	NumFrames
1	Glint	01-Dec-2014	12-Dec-2014	12	153	1,347,996
2	Nadir	13-Dec-2014	28-Dec-2014	16	228	1,857,425
3	Glint	04-Apr-2015	18-Apr-2015	15	198	1,650,609
4	Nadir	11-May-2015	21-May-2015	11	147	1,128,696

resolution Imaging Spectrometer (MODIS) aboard the *Aqua* platform provides complete collocation
 260 with the narrow swath ($\simeq 10$ km) of OCO-2. A variety of MODIS-*Aqua* and CALIOP products,
 subset to the OCO-2 ground tracks, are being produced by the A-Train Data Depot (ATDD) at the
 Goddard Earth Sciences Data and Information Services Center in collaboration with the Cooperative
 Institute for Research in the Atmosphere (CIRA) Data Processing Center (DPC) at Colorado State
 University (CSU). These products will prove useful for a variety of tasks such as spectral vicarious
 265 calibration of the instrument or detailed cloud and aerosol analysis.

The work presented here uses the collocated MODIS and CALIOP products to provide validation
 of the OCO-2 B7 cloud and aerosol screening for four sets of 16 day orbit repeat cycles in both nadir
 and glint viewing modes, as summarized in Table 2. No analysis of the special target viewing mode
 is reported in this work. See Wunch et al. (2015) for details.

270 Note that although OCO-2’s focal plane array detectors are equally divided into adjacent spatial
 subsets, providing eight footprints per measurement frame, this work reports only the results from
 footprint number four for clarity. No significant differences among the footprints were found in the
 statistical analysis.

Some brief analysis here on B7 cloud screening stats. Chris has compiled them into IDL save
 275 **files.** A quick analysis of the full OCO-2 B7 operational data set showed blah blah blah.

3.1 Collocation methodology

The ATDD generates collocated MODIS L1B and L2 atmospheric products for many of the satellites
 in the A-Train constellation using the algorithm described in (Savtchenko et al., 2008). The main dif-
 ference in the creation of the OCO-2 product relative to other A-Train sensors is the preparation of
 280 the reference track for ingest into the collocation algorithm. In the case of CloudSat, it is most conve-
 nient to use the Two-line Elements of the spacecraft to compute 15 minutes of CloudSat ground track
 for every MODIS 5-minute granule. However, the OCO-2 flight modes make this simple approach
 unattainable. Instead, the geolocation and time information of the central OCO-2 footprint must be
 extracted from an OCO-2 L1B science granule. Based on that time, an “OpenSearch” request is for-
 285 mulated and sent to the MODIS Processing System (LAADS). Upon acquiring the corresponding
 MODIS 5-minute granules (typically 9 per OCO-2 granule), a work order is logged with the God-

standard DAAC to drive MODIS granules one by one through the processing system. In most cases, part of the processing involves extrapolation of the OCO-2 track using the great arc model. The extrapolation is sufficiently accurate to extend the ground track of the OCO-2 footprint so that the resulting reference ground track fully transects the acquired MODIS granules. The resulting output are MODIS-like 5-minute HDF4 files that contain all the MODIS geolocation and science data for a given product within ± 50 km of the OCO-2 ground target.

The collocation of the OCO-2 soundings is performed at the DPC for MODIS L1B half-km radiances, 1- and 5-km cloud properties, and the 10-km aerosol product. As is done at the ATDD, the date and time information is extracted from the OCO-2 L1B files to determine the corresponding MODIS granules. Then any product-specific preprocessing is performed and a pixel-by-pixel matching to the OCO-2 geolocation is done. Finally, the 5-minute MODIS granules corresponding to a given OCO-2 granule are merged into a single file, such that there is a one-to-one file correspondence between the original OCO-2 granules and the collocated MODIS products.

The resultant DPC output HDF5 files contain both MODIS geolocation and science data within ± 50 km of the OCO-2 ground target, along with the geolocation and time information for OCO-2. Additionally, the files include the x and y MODIS pixel location for each match-up along with the origin filename. This configuration allows user customization of the match-up process, such as distance-dependent pixel searching.

The collocation procedure for CALIOP data is much the same, although **something, something....**

3.2 MODIS cloud mask

In this work we define a hybrid MODIS cloud mask by combining the standard cloud mask contained in the MYD06 cloud product (Ackerman et al., 1998; Frey et al., 2008) with the $1.3 \mu\text{m}$ cirrus reflectance value (Gao et al., 2002). A scene is defined as clear-sky when all MODIS pixels within 2 km of an OCO-2 sounding are labeled as *confident* or *probably* clear and the cirrus reflectance, \mathcal{R} , is less than 0.01. If any of these conditions are violated, then the true state of the scene is considered to be cloudy. In addition, we analyze only scenes with MODIS viewing zenith angle $< 30^\circ$, to avoid oblique lines of sight which can introduce errors (Maddux et al., 2010).

Tests were performed to ensure that the agreement between OCO-2 and MODIS cloud flags are not overly sensitive to the choice of the search radius, the cirrus reflectance or the sensor zenith angle. There is a small temporal discrepancy between the overpass time of OCO-2 and MODIS of about seven minutes, during which the geometrical and optical properties of clouds are subject to small changes and/or drifting in or out of the scene. However, errors in the validation procedure are mitigated by enforcing the 2 km radial search ($\simeq 12.5 \text{ km}^2$) when matching scenes, which provides about ten matching MODIS pixels per OCO-2 sounding.

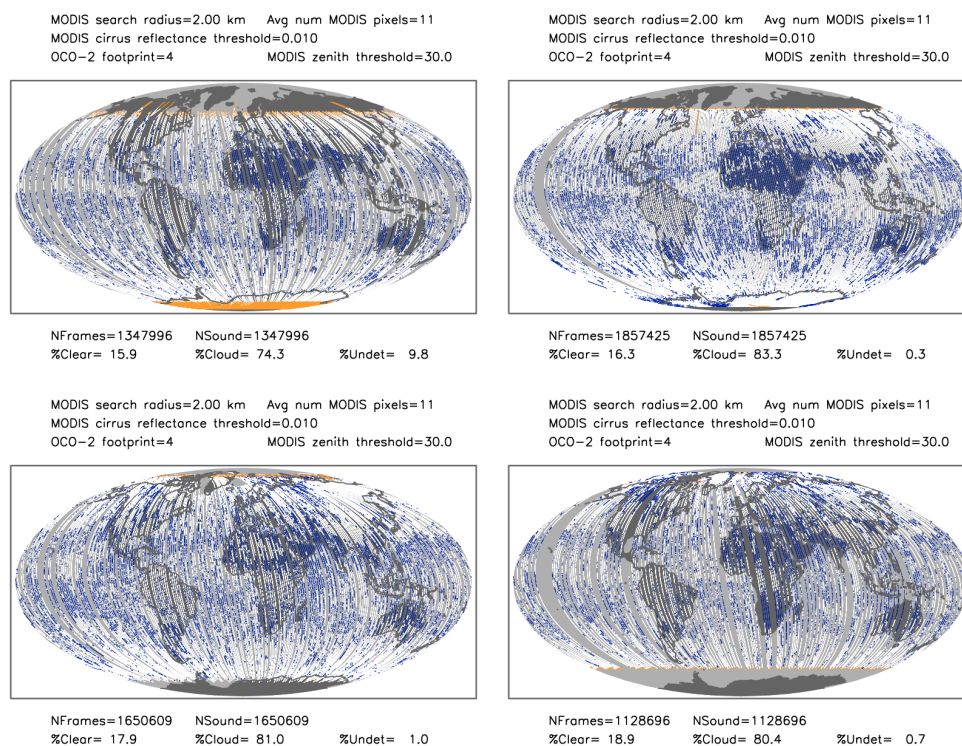


Figure 3. MODIS clear/cloudy results matched to the OCO-2 footprints for the winter glint (top left), the winter nadir (top right), spring glint (lower left) and spring nadir (lower right) orbits using the B7000 data set described in Section 3 (see Table 2). Cloudy points are first plotted in white, with clear points over-plotted in blue. The orange points represent sounding that had undetermined fill values in one or more data fields.

Figure 3 shows the OCO-2 orbit tracks for the four repeat cycles used in this analysis. The MODIS clear scenes are overlaid as blue points on the cloudy points, represented in white. Note that the individual points are not to proper scale as a single footprint size is only approximately 2 km cross track by 3 km along track. The maps are meant to indicate both the total global extent of the analyzed data set as well as the large spatially coherent cloudy and clear areas. Using the custom MODIS cloud mask described above, we find that the clear-sky fraction ranges from 16-19% over the four data sets. This result closely agrees with that reported in Figure 13 of Miller et al. (2007).

4 Validation of OCO-2 cloud screening

In this section we analyze the results of the comparison between the OCO-2 and MODIS cloud screening on the nadir and glint viewing repeat cycles that were introduced in Section 3. A brief, initial analysis of some spring-time CALIOP cloud and aerosol layer product collocated to OCO-2 is also presented.

Table 3. Example of a two dimensional contingency table.

	Predicted Clear	Predicted Cloudy
True Clear	True Positive (agree clear)	False Negative (disagree)
True Cloudy	False Positive (disagree)	True Negative (agree cloudy)

4.1 Contingency table analysis

335 Direct comparison of the OCO-2 preprocessor cloud flags to MODIS are analyzed using contingency
 tables (CT), which provide compact summary statistics for large data sets. This analysis follows that
 given in Taylor et al. (2012) on GOSAT data. A simple example of a CT is shown in Table 3 for
 illustrative purposes. For each sounding there are four possibilities; The diagonal elements indicate
 the number of collocated soundings that agree that the scene is clear or cloudy; true positives (TP)
 340 and true negatives (TN), respectively. The off-diagonal elements represent the number of scenes that
 disagree; false positives (FP) when MODIS indicates cloud but OCO-2 finds the scene to be clear.
 Finally, false negatives (FN) when MODIS indicates clear but OCO-2 cloud. For clarity of writing,
 the classification of scenes as either clear or cloudy by MODIS will be referred to as the “true” state,
 while scene classification by the OCO-2 preprocessors will be termed the “predicted” state.

345 Two statistics can be calculated from the CT values that help with the interpretation of the results.
 The accuracy (ACC) gives the fraction of scenes that are correctly predicted relative to the true state;

$$ACC = (N_{TP} + N_{TN})/N, \quad (1)$$

where N_{TP} , N_{TN} and N are the number of true positive, true negative and total scenes, respectively.

350 The positive predictive value (PPV) gives the percentage of the truly clear soundings that were
 predicted clear by the preprocessors;

$$PPV = N_{TP}/(N_{TP} + N_{FP}). \quad (2)$$

Prior to calculating the contingency table values on the test data set, it is best to optimize the ABP
 and IDP filter threshold values as is demonstrated in the following section.

355 4.2 Optimization of cloud screen thresholds

This section explores the optimization of the OCO-2 ABP and IDP parameters used to classify cloud
 contaminated scenes. Systematic variations in the preprocessor thresholds are expected to alter the
 model throughput (Γ = fraction of scenes passing the filter), accuracy (ACC) and positive predictive
 value (PPV). For any of the filters, a tightening (i.e., narrowing) of the thresholds generally yields
 360 an increase in ACC and PPV, while simultaneously decreasing Γ . This is an ideal situation if an

Table 4. Settings of the ABP and IDP filter thresholds used for the analysis presented in Section 4.3.

$\Delta p_{s,cld}$ (hPa)	χ^2 scale factor	r CO ₂ center point	r CO ₂ half-width	r H ₂ O center point	r H ₂ O half-width
25.0	5.0	1.01	0.03	0.98	0.08

aggressive cloud-clearing filter is desired. Conversely, if the design goal is to filter only the most grossly contaminated scenes while maximizing the throughput at the expensive of the accuracy, then the filtering thresholds should be set to relatively loose values.

The left column of Figure 4 shows contour plots of Γ , ACC and PPV for the spring-time nadir-land soundings as a function of the $\Delta p_{s,cld}$ and the χ^2 scaling factor, the two primary screening thresholds for the ABP algorithm. Optimal values, i.e., maximums, of the diagnostic parameters occur in the light shaded regions of each plot. The tradeoff in trying to maximize all three diagnostic parameters simultaneously is evident. In general, as ACC and PPV increase with tighter choices of $\Delta p_{s,cld}$ and χ^2 , Γ decreases. For this particular data set, setting $\Delta p_{s,cld}$ to 25 hPa and χ^2 scale factor to 5 allows $\Gamma \simeq 40\%$, ACC $\simeq 75\%$ and PPV $\simeq 50\%$.

The middle and right columns of Figure 4 show results from the sensitivity testing for the IDP r CO₂ and r H₂O halfwidth versus center point, respectively. To obtain Γ , ACC and PPV roughly equivalent to those described above for the ABP filters, settings for the r CO₂ and r H₂O given in Table 4 were selected. These values were then implemented in the cloud screen comparison to MODIS that will be detailed in Section 4.3. Note that the results for the glint-land and glint-water viewing scenarios looked similar to the nadir-land results presented here, allowing a single set of filters for the three viewing conditions. Slight differences were observed between the winter and spring data sets, indicating that slightly loose values of the thresholds should be employed to dissuade over filtering of the data.

4.3 OCO-2 cloud screen comparison to MODIS

After optimization of the thresholds, as discussed in Section 4.2, the contingency tables were calculated separately for each of the three viewing scenarios; nadir-land, glint-land and glint-water using the four repeat cycle data sets referenced in Table 2. The results of the CT analysis are displayed in Table 5. Note that for clarity the form of the contingency table has been shifted from two-by-two, as in Table 3, to a single row. However, the meaning of the results remains the same.

Overall the results are very encouraging. Using the selected thresholds, the throughput (Γ) is $\simeq 20\%$, with $\simeq 85\text{-}90\%$ agreement, i.e., accuracy, between MODIS and OCO-2 for all viewing conditions in the winter time data set. This in spite of the inherent errors in comparing cloud screening between satellite sensors with very different instrument characteristics and specifications.

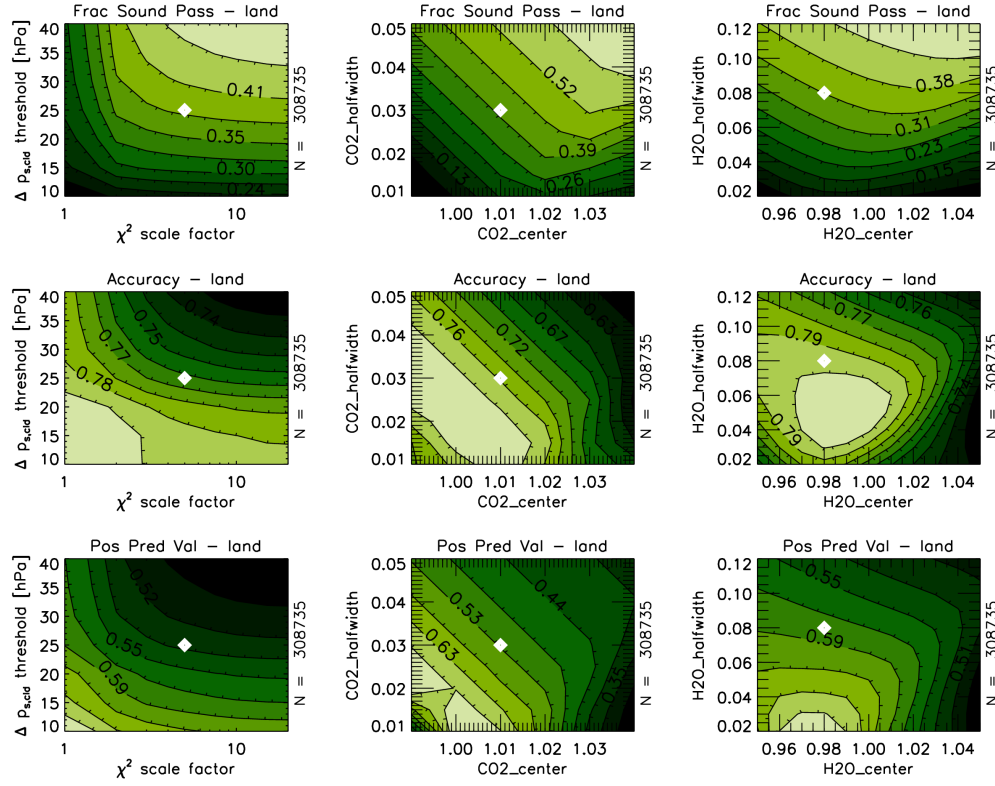


Figure 4. Contour plots showing the fraction of soundings passing (top row), accuracy (middle row) and positive predictive value (bottom row) for variations in the ABP $\Delta p_{s,cl\Delta}$ versus χ^2 scale factor (left column), IDP r CO₂ (middle column) and r H₂O (right column). All results are for nadir-land scenes for the spring-time data set. The white diamond represents the threshold settings adopted for the analysis presented in this work.

Table 5. Contingency tables for the comparison of the OCO-2 preprocessors to MODIS cloud screening. Results are shown for the three main viewing scenarios for both the winter and spring-time data sets.

Viewing	True clear atmospheres					True cloudy atmospheres					Γ	ACC	PPV
	Total cases	Predicted clear	Predicted cloudy			Total cases	Predicted clear	Predicted cloudy					
		N _{TP}	TPR	N _{FN}	FNR		N _{FP}	FPR	N _{TN}	TNR			
Winter													
Nadir-land	105686	69507	65.8%	36179	34.2%	548475	58480	10.7%	489995	89.3%	19.6%	85.5%	54.3%
Glint-land	71416	52515	73.5%	18901	26.5%	301577	22840	7.6%	278737	92.4%	20.2%	88.8%	69.7%
Glint-water	137656	107290	77.9%	30366	22.1%	692783	76435	11.0%	616348	89.0%	22.1%	87.1%	58.4%
Spring													
Nadir-land	81013	50737	62.6%	30276	37.4%	226006	25152	11.1%	200854	88.9%	24.7%	81.9%	66.9%
Glint-land	91854	56021	61.0%	35833	39.0%	353099	37091	10.5%	316008	89.5%	20.9%	83.6%	60.2%
Glint-water	195840	140381	71.7%	55459	28.3%	959763	85614	8.9%	874149	91.1%	19.6%	87.8%	62.1%

Table 6. Contingency table analysis similar to Table 5, but for all eight footprints for the winter-time glint-water data set.

Footprint	True clear atmospheres					True cloudy atmospheres					Γ	ACC	PPV
	Total cases	Predicted clear		Predicted cloudy		Total cases	Predicted clear		Predicted cloudy				
		N _{TP}	TPR	N _{FN}	FNR		N _{FP}	FPR	N _{TN}	TNR			
1	138341	106018	76.6%	32323	23.4%	691724	73285	10.6%	618439	89.4%	21.6%	87.3%	59.1%
2	138022	106936	77.5%	31086	22.5%	692075	73900	10.7%	618175	89.3%	21.8%	87.4%	59.1%
3	137802	107313	77.9%	30489	22.1%	692445	75780	10.9%	616665	89.1%	22.1%	87.2%	58.6%
4	137656	107290	77.9%	30366	22.1%	692783	76435	11.0%	616348	89.0%	22.1%	87.1%	58.4%
5	138034	106860	77.4%	31174	22.6%	692470	74079	10.7%	618391	89.3%	21.8%	87.3%	59.1%
6	138038	107815	78.1%	30223	21.9%	692580	75048	10.8%	617532	89.2%	22.0%	87.3%	59.0%
7	138179	107837	78.0%	30342	22.0%	692513	75933	11.0%	616580	89.0%	22.1%	87.2%	58.7%
8	138282	106833	77.3%	31449	22.7%	692481	73568	10.6%	618913	89.4%	21.7%	87.4%	59.2%

The positive predictive values (PPV) range from 54% to 70% for the winter time data, meaning that for every ten sounding passed as clear by the combined ABP + IDP, MODIS agrees that 5 to 7 of them are clear. The somewhat low values of PPV are largely driven by the dominance of cloudy over clear scenes. Although OCO-2 successfully flags $\simeq 90\%$ of the true cloudy scenes, the remaining 10% is a large fraction ($\simeq 40\%$) of the total passing as clear. While on the one hand this would appear to be a rather large ratio of contaminated scenes to be passing to the L2 X_{CO_2} retrieval algorithm, recall that the actual operational sounding selection will be performed via an advanced machine learning technique (Mandrake et al., 2013) that uses as inputs, output from ABP and IDP, as well as from L1B. Also note that these statistics are in reference to MODIS as truth, an assumption that is not void of uncertainties. Furthermore, the cloud screen threshold values were selected here to be relatively loose. As was discussed in Section 4.2, tighter thresholds generally increase PPV, but at the expense of Γ .

Another phenomena of note from the contingency table analysis is the high agreement in cloudy scenes between the two sensors (TNR $\simeq 8$ -12 times larger than FPR) versus the much lower agreement when MODIS passes the scene as clear (TPR only $\simeq 2$ -3 times larger than FPR). This seems to imply that OCO-2 is much more sensitive to contamination by cloud and aerosol than is MODIS, and the effect seems to be strongest over land.

An investigation of the eight footprints individually via CT statistics revealed no strong dependence. For completeness, the results for the winter time glint-water case are shown in Table 6. The range of variability for all of the CT values is always well under 2% and is generally closer to 1%.

It is critical to avoid spatial sampling biases in the measurement of X_{CO_2} , as these can yield serious errors in flux inversion estimates, e.g., Liu et al. (2014). To this end, it is desirable for the OCO-2 cloud screening algorithm to properly pass a reasonable fraction of true clear-sky scenes in persistently cloudy regions such as the convective tropics.

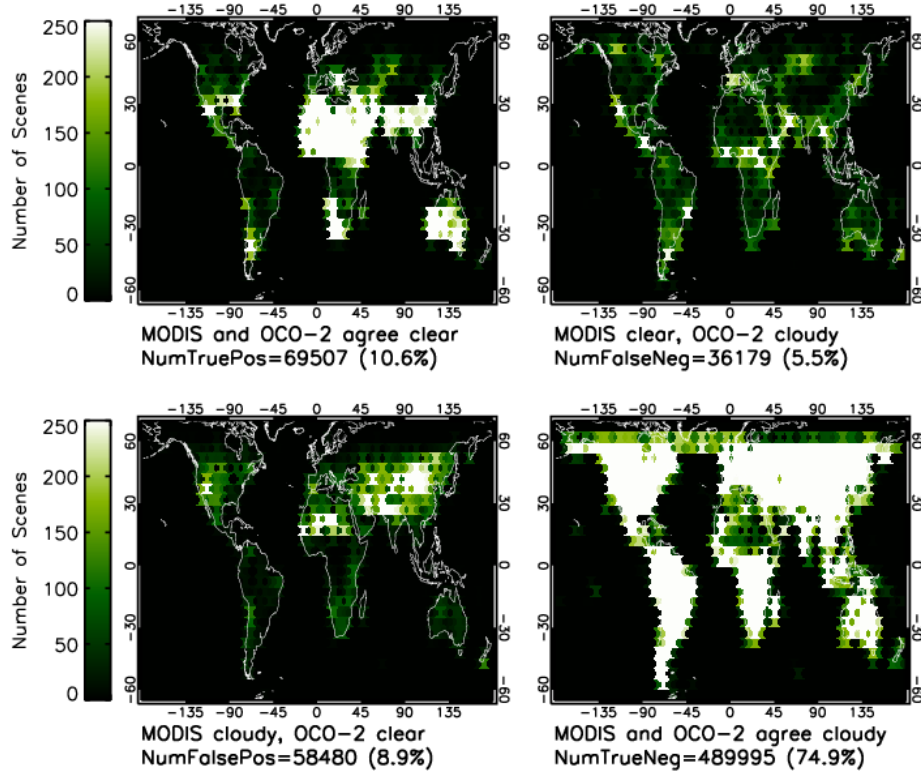


Figure 5. Sounding density in each of the four contingency table classification bins using the combined ABP + IDP cloud filters for the winter time nadir-land viewing data. Data are binned on a 5° by 5° lat/lon grid and shown from 60° S to 70° N latitude. Correct classifications are true positives (upper left), true negatives (lower right), while incorrect classifications are given by false negatives (upper right) and false positives (lower left).

To access the spatial distributions of the contingency table values, soundings were gridded into 5° latitude by 5° longitude boxes and plotted as the sounding density per box, as seen in Figure 5 for the winter time nadir-land viewing case. One of the key conclusions from these maps is that a large number of clear soundings are found over climatologically dry land regions like the Sahara and interior Australia, as expected (upper left panel). Most land regions contain a large number of soundings determined to be cloudy by both sensors (lower right panel), while the midlatitude and tropical land regions have very mixed results, with substantial numbers of soundings in each of the four classification types. Note that the maps have been truncated to exclude data poleward of 60° S and 70° N. Soundings in the polar regions are typically problematic for the L2 X_{CO_2} retrieval algorithm due to snow and ice covered surfaces, in conjunction with very high SZA.

Further analysis was performed by calculating the fraction of the scenes per lat/lon grid box for each of the four CT bins, as shown in Figure 6. For any given lat/lon bin, the fractional values from all four of the panels by definition sums to 1.0. Over the dry land regions we see the highest fraction

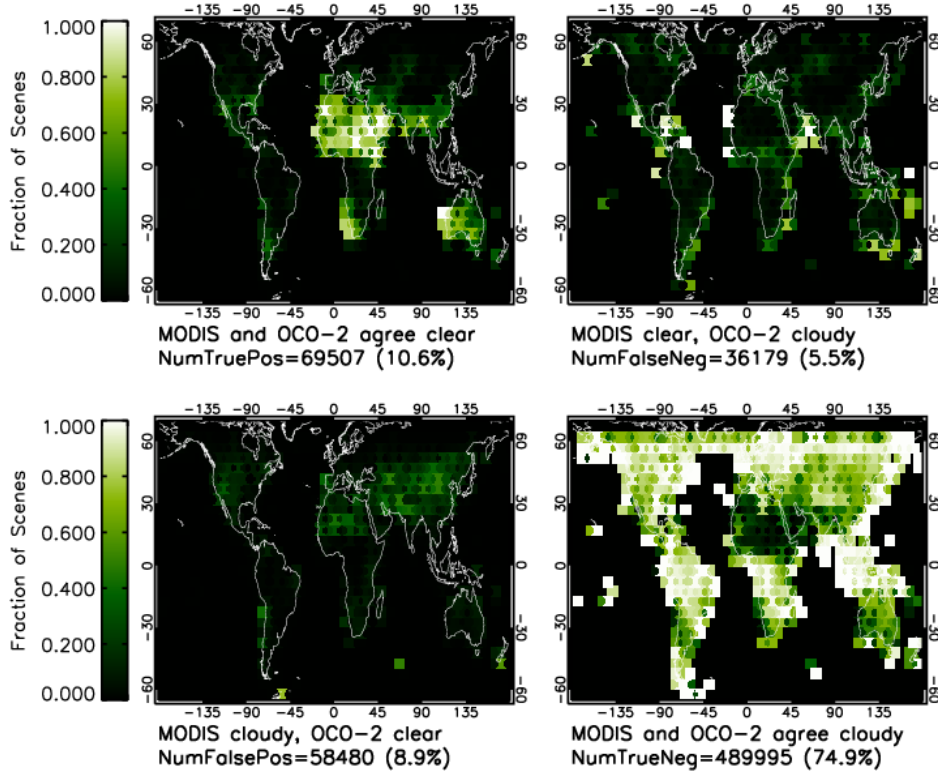


Figure 6. The fraction of the total number of scenes in each lat/lon bin for each of the four contingency table classifications using the combined ABP + IDP cloud filters for the winter time nadir-land viewing data. Data are binned on a 5° by 5° lat/lon grid and shown from 60° S to 70° N latitude. Correct classifications are true positives (upper left), true negatives (lower right), while incorrect classifications are given by false negatives (upper right) and false positives (lower left).

of agreement in clear scenes between OCO-2 and MODIS, while large areas of the high northern latitudes and convective tropical land regions (Amazonia, Congo, Indonesia, etc) are almost exclusively classified as cloudy by both instruments. In the northern hemisphere mid-latitudes there is a persistent belt spanning roughly 25° N to 55° N that contains 10-40% false positive scenes (MODIS cloudy, OCO-2 clear), persisting across all longitudes. Finally, there is a very spatially independent distribution of false negative scenes (MODIS clear, OCO-2 cloudy) with a fractional value of 10 to 30% with some higher fractions sprinkled across the globe. Of note is that a disproportionate amount of these FN soundings occur along coastlines.

Figure 7 and Figure 8 show the sounding density and contingency table fraction maps for the nadir-land spring time data set. Some differences in the spatial distributions relative to the winter time data are evident such as reduced fractional agreement of clear scenes and a significant increase in the number of FN soundings (MODIS clear, OCO-2 cloudy) over the Sahara region. It is likely

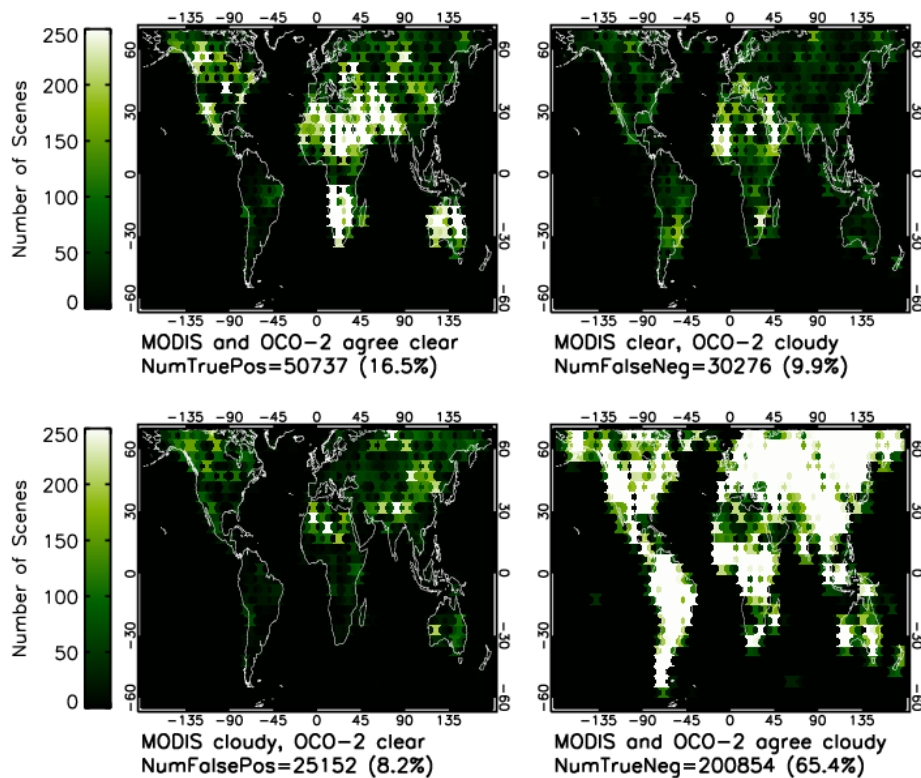


Figure 7. Same as Figure 5, but for nadir-land spring time.

that this is a misclassification of dust by MODIS. **need to try MODIS dust filter!** In general though, the patterns look much the same, indicating that there is unlikely to be strong sampling biases in the cloud screened data set as a function of season.

To show more clearly the spatial distribution of the agreement between MODIS and OCO-2, the calculated accuracies in 5° by 5° lat/lon bins for the winter and spring data sets are shown in Figure 9 for all three viewing conditions; nadir-land, glint-land and glint-water.

For nadir-land (top row) the agreement is best ($>90\%$) in the high latitude regions as well as near the sub-solar point (about 23° S for the winter time data set). The worst agreement is found in a large contiguous regions spanning the Sahara, the Saudi Arabian peninsula, the Himalayan region and east through China. The effect strengthens from the winter to the spring time. The agreement also substantially worsens over northern North America over this time period. The reason for this is rooted in the relatively low agreement rate for the scenes MODIS deems clear, coupled with a large number of MODIS clear soundings over the described region. It is possible that MODIS is misclassifying as clear many scenes containing desert dust (Sahara and Arabia), those with high vertical relief (Himalayas and Rocky Mountains) and thick aerosols (China). These very same scenes are then being identified as cloud and/or aerosol contaminated by the ABP and IDP filters. Table ??

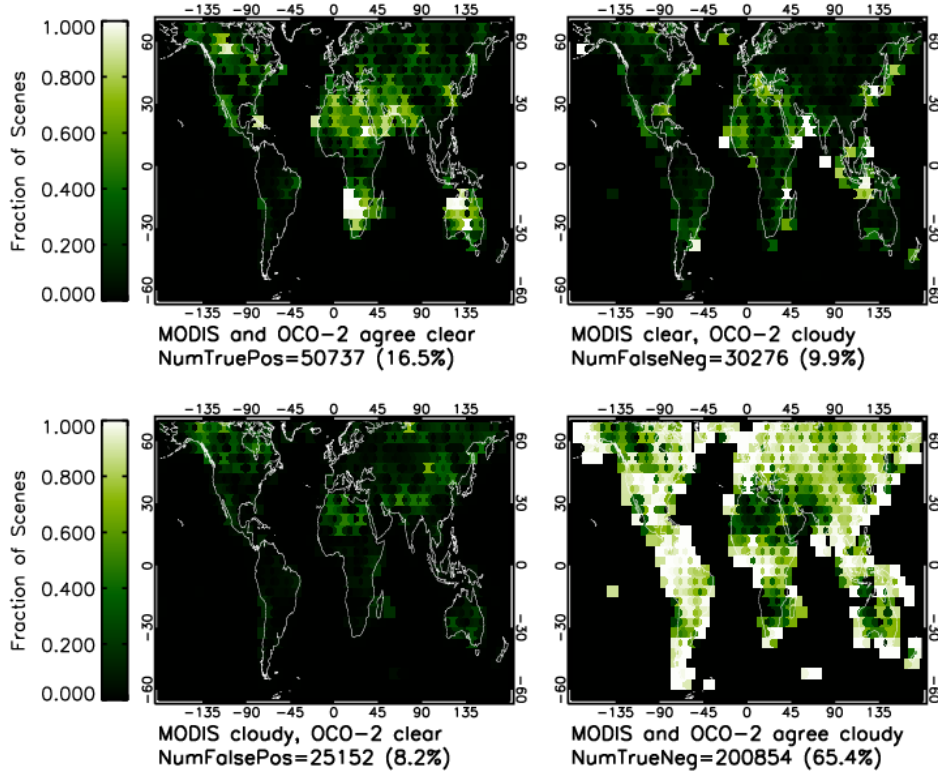


Figure 8. Same as Figure 6, but for nadir-land spring time.

455 indicates that there was an increase in both Γ and PPV from the winter to the spring, while the ACC decreased substantially for nadir-land.

The middle row of Figure 9 shows the percent agreement for the glint-land viewing data. Overall, the features are very similar to those for the nadir-land, although the region of strong agreement in the high latitude northern hemisphere extends further south, possibly due to increased SNR in glint
 460 mode. Furthermore, the disagreement over the Sahara and Arabia regions seems to be less severe than it was for nadir-land. Recall from Table ?? that the ACC and PPV were substantially worse from winter to spring in the global mean.

Finally, the bottom row of Figure 9 shows the accuracy for the glint-water viewing data. The best agreement is again found at the higher latitudes (both north and south), while the worst agreement
 465 occurs in the tropical belt. This is again consistent with the fact that the OCO-2 and MODIS screening agree well for cloudy scenes, but are less in agreement for scenes MODIS says are clear. Note that the ACC and PPV values reported in Table 5 do not vary much from the winter to the spring data set for glint-water viewing.

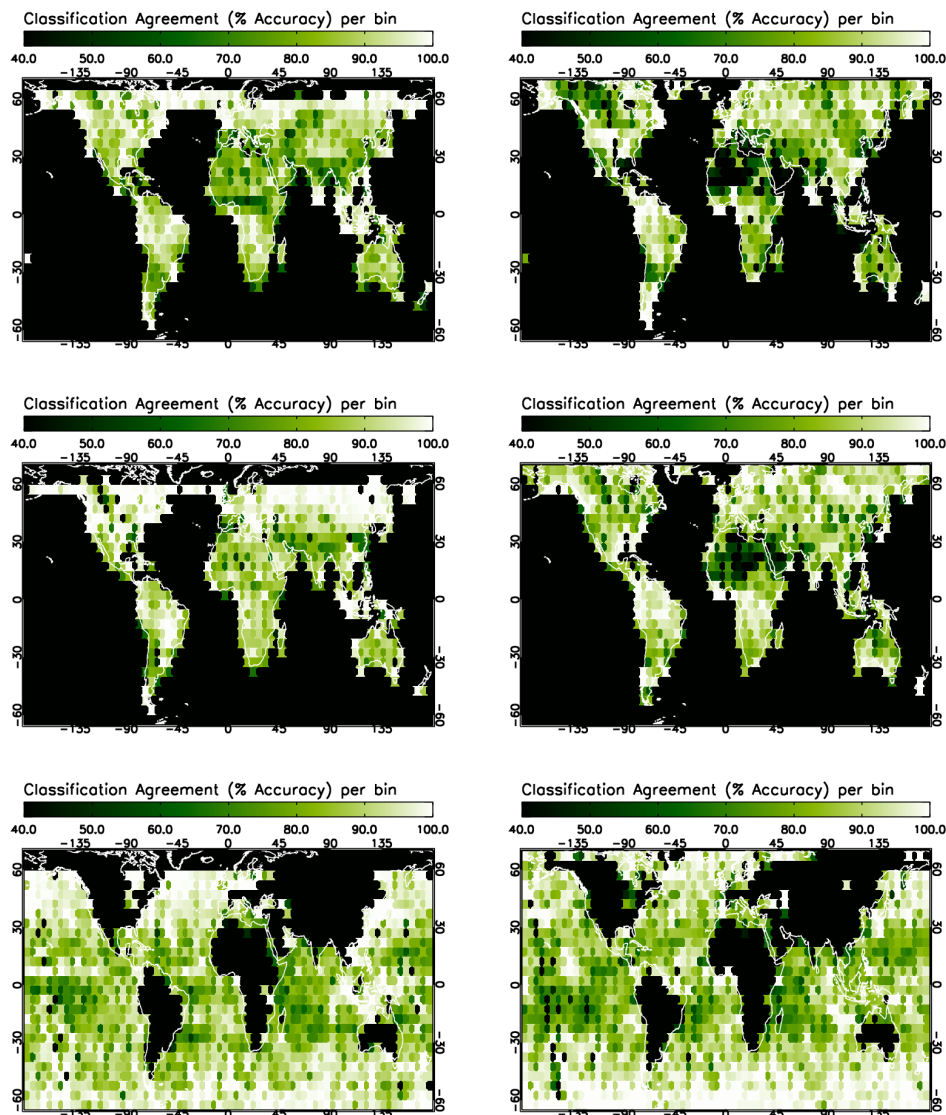


Figure 9. The accuracy of the cloud classification agreement between OCO-2 and MODIS for the winter (left column) and spring (right column) for nadir-land (top), glint-land (middle) and glint-water (bottom) data sets.

4.4 OCO-2 cloud screen comparison to CALIPSO

470 Figure 10 shows analysis of the spring nadir-land data set from the OCO-2 CALIOP matchup. The bar histogram on the left panel (all scenes in green and clear scenes in blue) indicates a reasonable distribution of TOD, ranging from 0.01 to 10, at which point the CALIOP instrument saturates. For this particular data set there is a large spike in the number of scenes with $TOD \simeq 2-3$, which seems to be a real, robust geophysical signature.

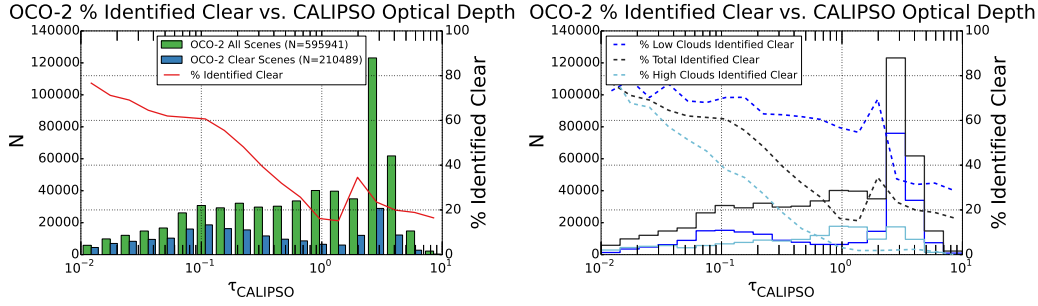


Figure 10. Histograms of the number of OCO-2 scenes identified as clear (blue) and total number of scenes (green) versus the CALIPSO optical depth (left panel). The percent of scenes identified as clear by OCO-2 (ratio of blue to green) are shown by the red trace. The right panel shows the same data subsetted by high and low clouds (see text for definitions). Note that the black dashed line, representing the total scenes identified as clear, is by definition equivalent to the red curve in the left panel.

475 The red trace in the plot shows that the fraction of scenes that the ABP flags as cloudy is heavily correlated with the CALIOP TOD, as expected. That is, there is a decrease in the percent identified as clear with increasing TOD. Note that here only the operational ABP cloud flag is used, i.e., $\Delta p_{s, cld} < \pm 30$ hPa (the χ^2 filter is essentially disabled in the current code by setting an extraordinarily high value for the threshold). For TOD $\simeq 0.1$, about 60% of the scenes are identified as clear (40% flagged as cloudy) by the ABP, while between 15 and 35% are passed when TOD > 1 .

Recall from the study on simulated data that was presented in Section 2.3, it was found that ABP tends to work well in identifying high cloud layers but is more prone to error when the cloud scattering height is near to the surface. This effect is demonstrated for the first time on real data in the left panel of Figure 10 which shows the same data subsetted by cloud height, where high (low) is defined as cases where 95% of the TOD resides in the top 40% (bottom 30%) of the atmosphere. It is evident that the ABP algorithm is passing a very large fraction of the scenes containing low clouds (60% at TOD = 1), while it is very aggressive at screening high clouds (less than 5% with TOD = 1 pass).

To access the change in the screening when combining the two cloud screening algorithms, a similar analysis was performed, but this time using the ABP and IDP filter thresholds set to the values reported in Table 4. The results are shown in Figure 11, which indicates the drastically reduced throughput. For TOD $\simeq 0.1$, about 30% of the scenes are identified as clear (half of what was found for ABP alone), while less than 10% pass when TOD > 1 . The right panel demonstrates that the combination of the two filters partially mitigates the issue of passing thick low clouds, as only about 20% of the scenes with TOD = 1 are passing (compared to 60% for the ABP alone).

To further access scenes with relatively high optical thicknesses that are erroneously passing the ABP cloud flag, a subset of the data was created to include only those soundings with CALIOP TOD > 1 . The effective cloud top pressure (p_c) was calculated by integrating the optical depth vertically

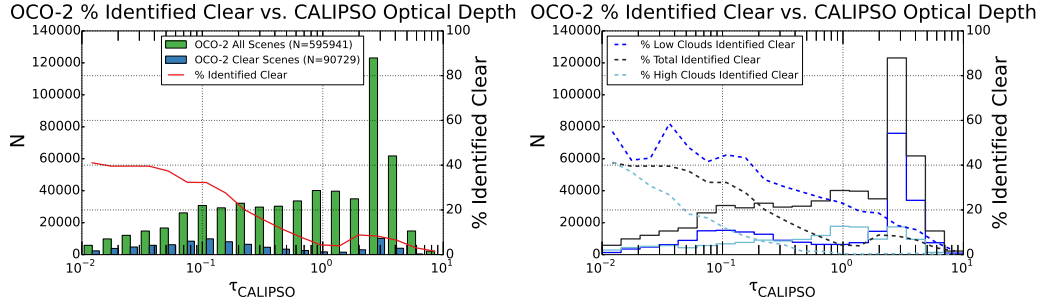


Figure 11. Same as Figure 10, but using the combined ABP and IDP preprocessors. The much more aggressive filtering (reduction in throughput) and the identification of optically thick, low clouds is apparent.

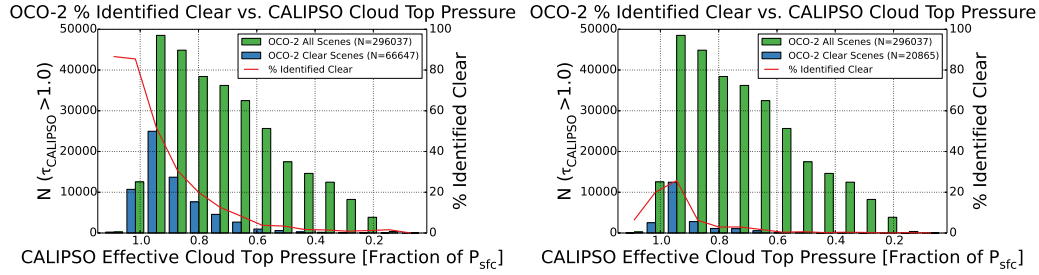


Figure 12. Histograms of the number of OCO-2 scenes identified as clear (blue) and total number of scenes (green) versus the CALIPSO optical depth (left panel). The percent of scenes identified as clear by OCO-2 (ratio of blue to green) are shown by the red trace. The right panel shows the histograms versus the calculated effective cloud top pressure for only scenes with CALIPSO TOD > 1.

down through the atmosphere (starting at the top) until $TOD > 1$ was achieved. The pressure value of that layer was then assumed to represent p_c . This value was then normalized by the ECMWF model surface pressure, represented by \bar{p}_c . Thus p_c near the surface corresponds to \bar{p}_c values near 1, while p_c higher in the atmosphere are represented by \bar{p}_c values near 0.

The results from this analysis are shown in Figure 12, which again shows the frequency distribution (bar plots) and the fraction of scenes identified as clear (red trace). The left plot uses only the ABP cloud flag, while the right plot utilizes the combined filter. The drastic difference between the two cases is evident. For ABP alone, at $\bar{p}_c \approx 0.95$, more than 50% of the scenes are passing, while for the combined filter only about 20% of these low, optically thick scenes pass. This is confirmation that the ABP, which relies on photon path length modification to detect cloud, is unable to discern cloud near the surface, even when the optical thickness is large. Conversely, ABP is adept at picking out very thin scattering layers when they are located high in the atmosphere due to the strong photon path length modification. But the ABP and IDP filters in unison, provide a very robust cloud screening that catches most of the optically thick, low level clouds.

5 Conclusions

In summary, the OCO-2 preprocessors are performing well and meeting the need for prescreening the data before further processing in X_{CO_2} L2 retrieval algorithm. We have demonstrated that the ABP and IDP algorithms can be sufficiently tuned to pass $\simeq 20\%$ of the data while maintaining overall agreement of 85-90% with the MODIS cloud mask.

The spatial distribution of passed soundings is of critical concern to avoid spatial sampling biases which can lead to errors in source/sink inversion models. We found that the preprocessors are passing soundings over all portions of the globe, although higher latitudes and higher solar zenith angles tend to be problematic, due to snow and ice covered surfaces and lower signal to noise ratios, respectively, both of which make accurate cloud identification and X_{CO_2} retrievals difficult.

In this work we also demonstrated the strength of the ABP in identifying high, thin cloud layers via a simulated study and a comparison to collocated CALIOP data. In addition, we were able to access the ABP's weakness for passing clouds low to the surface, even those with total optical depths well above what can be accurately modeled in the L2 retrieval algorithm. However, the combination of the ABP with the IDP mitigates the number of the low, thick clouds that are passing.

Detailed studies uncovered no significant time-dependent or footprint-dependent features in the OCO-2 prescreening data. The current B7 operational data set is therefore providing a robust cloud-cleared data set for use in scientific investigations of global CO_2 .

Acknowledgements. The CSU contribution to this work was supported by JPL contracts 1380533 and 5319950. A portion of the research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

We would like to acknowledge the hard work of many individuals on the OCO-2 algorithm, data processing and science teams. Although they did not contribute directly to this manuscript, their efforts made this work possible. The list (in alphabetical order) includes (but is not limited to); Charlie Avis, Lars Chapsky, Lan Dang, Brendan Fisher, Robert Granat, Richard Lee, Lukas Mandrake, James McDuffie, Vijay Natraj, Greg Osterman, Fabian Oyafuso, Vivienne Payne, Randy Pollock, Rob Rosenberg, Mike Smyth, Paul Wennberg, Debra Wunch and Jia Zong.

In addition we thank the computer technical support staff at CSU; Natalie Tourville and Michael Hiatt.

References

- Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., and Gumley, L. E.: Discriminating clear sky from clouds with MODIS, *J. Geophys. Res.*, 103, D07206, doi:10.1029/1998.
- Boesch, H., Brown, L., Castano, R., Christi, M., Connor, B., Crisp, D., Eldering, A., Fisher, B., Frankenberg, C., Gunson, M., Granat, R., McDuffie, J., Miller, C., Natraj, V., O'Brien, D., O'Dell, C., Osterman, G., Oyafuso, F., Payne, V., Polonsky, I., Smyth, M., Spurr, R., Thompson, D., and Toon, G.: Orbiting Carbon Observatory (OCO)-2 Level 2 Full Physics Algorithm Theoretical Basis Document, Algorithm Theoretical Basis Document OCO D-55207, Jet Propulsion Laboratory, 2015.
- Crisp, D. and Johnson, C.: The Orbiting Carbon Observatory mission, *J. Atmos. Sci.*, 56, 193–197, doi:10.1016/j.actaastro.2004.09.032, 2005.
- Crisp, D., Atlas, R., Breon, F.-M., Brown, L. R., Burrows, J., Ciais, P., Connor, B. J., Doney, S. C., Fung, I. Y., Jacob, D. J., Miller, C. E., O'Brien, D., Pawson, S., Randerson, J. T., Rayner, P., Salawitch, R. J., Sander, S. P., Sen, B., Stephens, G. L., Tans, P. P., Toon, G. C., Wennberg, P. O., Wofsy, S. C., Yung, Y. L., Kuang, Z., Chudasama, B., Sprague, G., Weiss, B., Pollock, R., Kenyon, D., and Schroll, S.: The Orbiting Carbon Observatory (OCO) mission, *Adv. Space Res.*, 34, 700–709, doi:10.1016/j.asr.2003.08.062, 2004.
- Day, J. O., O'Dell, C. W., Pollock, R., Bruegge, C. J., Rider, D., Crisp, D., and Miller, C. E.: Preflight spectral calibration of the Orbiting Carbon Observatory, *IEEE Trans. Geosci. Remote Sens.*, 49, 2793–2801, doi:10.1109/TGRS.2011.2107745, 2011.
- Frankenberg, C.: OCO-2 IMAP-DOAS preprocessor, Algorithm Theoretical Basis Document, Jet Propulsion Laboratory, 2014.
- Frankenberg, C., Platt, U., and Wagner, T.: Iterative maximum a posteriori (IMAP)-DOAS for retrieval of strongly absorbing trace gases: Model studies for CH₄ and CO₂ retrieval from near infrared spectra of SCIAMACHY onboard ENVISAT, *Atmos. Chem. Phys.*, 5, 9–22, 2005.
- Frey, R. A., Ackerman, S. A., Liu, Y., Strabala, K. I., Zhang, H., Key, J. R., and Wang, X.: Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5, *J. Atmos. Oceanic Technol.*, 25, 1057–1072, doi:10.1175/2008JTECHA1052.1, 2008.
- Gao, B., Yang, P., Han, W., Li, R.-R., and Wiscombe, W.: An algorithm using visible and 1.375- μ m channels to retrieve cirrus cloud reflectances from aircraft and satellite data, *IEEE Trans. Geosci. Remote Sens.*, 40, 1659–1688, doi:10.11, 2002.
- Ishida, H. and Nakajima, T. Y.: Development of an unbiased cloud detection algorithm for a spaceborne multi-spectral imager, *J. Geophys. Res.*, 114, doi:10.1029/2008JD010710, 2009.
- Ishida, H., Nakajima, T. Y., Yokota, T., Kikuchi, N., and Watanabe, H.: Investigation of GOSAT TANSO-CAI cloud screening ability through an intersatellite comparison, *J. Appl. Meteorol. Clim.*, 50, 1571–1586, doi:10.1175/2011JAMC2672.1, 2011.
- Kuze, A., Suto, H., Nakajima, M., and Hamazaki, T.: Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring, *Appl. Opt.*, 48, 6716–6733, 2009.
- Lee, R.: Preflight Spectral Calibration of the Orbiting Carbon Observatory-2, in preparation, x, 2015.
- Liu, J., K., B., Lee, M., Henze, D., Bousserez, N., Brix, H., Collatz, G., Menemenlis, D., Ott, L., Pawson, S., Jones, D., and Nassar, R.: Carbon monitoring system flux estimation and attribution: impact of

ACOS-GOSAT XCO₂ sampling on the inference of terrestrial biospheric sources and sinks, *Tellus B*, 66, doi:10.3402/tellusb.v66.22486, 2014.

Maddux, B., Ackerman, S., and Platnick, S.: Viewing geometry dependencies in MODIS cloud products, *J. Atmos. Oceanic Technol.*, 27, 1519–1528, doi:10.1175/2010JTECHA1432.1, 2010.

585 Mandrake, L., Frankenberg, C., O'Dell, C., Osterman, G., Wennberg, P., and Wunch, D.: Semi-autonomous sounding selection for OCO-2, *Atmos. Meas. Tech.*, 6, 2851–2864, doi:10.5194/amt-6-2851-2013, 2013.

Mandrake, L., Doran, G., O'Dell, C., Wunch, D., Wennberg, P., Eldering, A., Fischer, B., Osterman, G., Frankenberg, C., and Gunson, M.: OCO-2's Quality Estimation and Data Filtering Product: A Complete Description of the DOGO Warn Levels from Construction to Function, in preparation, x, 2015.

590 Miller, C., Crisp, D., DeCola, P., Olsen, S., Randerson, J., Michalak, A., Alkhaled, A., Rayner, P., Jacob, D., Suntharalingam, P., Jones, D., Denning, A., Nicholls, M., Doney, S., Pawson, S., Boesch, H., Connor, B., Fung, I., O'Brien, D., Salawitch, R., Sander, S., Sen, B., Tans, P., Toon, G., Wennberg, P., Wofsy, S., Yung, Y., and Law, R.: Precision requirements for space-based X_{CO_2} data, *J. Geophys. Res.*, 112, doi:10.1029/2006JD007659, 2007.

595 O'Brien, D. M., Polonsky, I., O'Dell, C., and Carheden, A.: Orbiting Carbon Observatory (OCO), algorithm theoretical basis document: The OCO simulator, Technical report ISSN 0737-5352-85, Cooperative Institute for Research in the Atmosphere, Colorado State University, 2009.

O'Dell, C., Taylor, T. E., and Eldering, A.: OCO-2 Oxygen-A Band Cloud Screening Algorithm (ABO2), Algorithm Theoretical Basis Document, Jet Propulsion Laboratory, 2014.

600 O'Dell, C. W., Day, J. O., Pollock, R., Bruegge, C. J., O'Brien, D. M., Castano, R., Tkatcheva, I., Miller, C. E., and Crisp, D.: Preflight radiometric calibration of the Orbiting Carbon Observatory, *IEEE Trans. Geosci. Remote Sens.*, 49, 2438–2447, doi:10.1109/TGRS.2010.2090887, 2011.

O'Dell, C. W., Connor, B., Bösch, H., O'Brien, D., Frankenberg, C., Castano, R., Christi, M., Crisp, D., Eldering, A., Fisher, B., Gunson, M., McDuffie, J., Miller, C. E., Natraj, V., Oyafuso, F., Polonsky, I., Smyth, 605 M., Taylor, T., Toon, G. C., Wennberg, P. O., and Wunch, D.: The ACOS CO₂ retrieval algorithm—Part 1: Description and validation against synthetic observations, *Atmos. Meas. Tech.*, 5, 99–121, doi:10.5194/amt-5-99-2012, 2012.

Rosenberg, R., Maxwell, S., Johnson, B. C., Chapsky, L., Lee, R., and Pollock, R.: Preflight Radiometric Calibration of the Orbiting Carbon Observatory-2, in preparation, x, 2015.

610 Savtchenko, A., Kummerow, A., Smith, P., Gopalan, A., Kempler, S., and Leptoukh, G.: A-Train Data Depot - Bringing atmospheric measurements together, *IEEE Trans. Geosci. Remote Sens.*, 46, 2788–2795, 2008.

Seidel, F. and Popp, C.: Critical surface albedo and its implications to aerosol remote sensing, *Atmos. Meas. Tech.*, 5, 1653–1665, doi:10.5194/amt-5-1653-2012, 2012.

Stephens, G. L., Vane, D. G., Boain, R. J., and Kenneth Sassen, G. G. M., Wang, Z., Illingworth, A. J., O'Connor, 615 E. J., Rossow, W. B., Durden, S. L., Miller, S. D., Austin, R. T., Benedetti, A., and Mitrescu, C.: The Cloudsat mission and the A-train: A new dimension of space-based observations of clouds and precipitation, *Bulletin of the American Meteorological Society*, December, 1771–1790, 2002.

Taylor, T. E. and O'Brien, D. M.: A neural network cloud screening algorithm. Part I: a synthetic case over land surfaces using micro-windows in O₂ and CO₂ near infrared absorption bands with nadir viewing, *J. Appl.*

620 *Rem. Sens.*, 3, 033 548, doi:10.1117/1.3239515, 2009.

Taylor, T. E., O'Dell, C. W., O'Brien, D. M., Kikuchi, N., Yokota, T., Nakajima, T. Y., Ishida, H., Crisp, D., and Nakajima, T.: Comparison of cloud-screening methods applied to GOSAT near-infrared spectra, *IEEE Trans. Geosci. Remote Sens.*, 50, 295–309, doi:10.1109/TGRS.2011.2160270, 2012.

Wunch, D., Toon, G., Wennberg, P., Wofsy, S., Stephens, B., Fischer, M., Uchino, O., Abshire, J., Bernath, P.,
 625 Biraud, S., Blavier, J., Boone, C., Bowman, K., Browell, E., Campos, T., Connor, B., Daube, B., Deutscher, N., Diao, M., Elkins, J., Gerbig, C., Gottlieb, W., Griffith, D., Hurst, D., Jimenez, R., Keppel-Aleks, G., Kort, E., Macatangay, R., Machida, T., Matsueda, H., Moore, F., Morino, I., Park, S., Robinson, J., Roehl, C., Sawa, Y., Sherlock, V., Sweeney, C., Tanaka, T., and Zondlo, M.: Calibration of the Total Carbon Column Observing Network using aircraft profile data, *Atmos. Meas. Tech.*, 3, 1351–1362, doi:10.5194/amt-3-1351-2010, 2010.

Wunch, D., Wennberg, P., Toon, G., Connor, B., Fisher, B., Osterman, G., Frankenberg, C., Mandrake, L., O'Dell, C., P.Ahonen, Biraud, S., Castano, R., Cressie, N., Crisp, D., Deutscher, N., Eldering, A., Fisher, M., Griffith, D., Heikkinen, M. G. P., Kyro, G. K.-A. E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C., Morino, I., Notholt, J., Oyafuso, F., Rettinger, M., Robinson, J., Roehl, C.,
 635 Salawitch, R., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D., Uchino, O., Warneke, T., and Wofsy, S.: A method for evaluating bias in global measurements of CO₂ total columns from space, *Atmos. Chem. Phys.*, 11, 12 317–12 337, doi:10.5194/acp-11-12317-2011, 2011.

Wunch, D., Wennberg, P., Toon, G., Connor, B., Fisher, B., Osterman, G., Frankenberg, C., Mandrake, L., O'Dell, C., P.Ahonen, Biraud, S., Castano, R., Cressie, N., Crisp, D., Deutscher, N., Eldering, A., Fisher, M., Griffith, D., Heikkinen, M. G. P., Kyro, G. K.-A. E., Lindenmaier, R., Macatangay, R., Mendonca, J., Messerschmidt, J., Miller, C., Morino, I., Notholt, J., Oyafuso, F., Rettinger, M., Robinson, J., Roehl, C.,
 640 Salawitch, R., Sherlock, V., Strong, K., Sussmann, R., Tanaka, T., Thompson, D., Uchino, O., Warneke, T., and Wofsy, S.: Early analysis of OCO-2 target mode data against TCCON, in preparation, x, 2015.

Yoshida, Y., Eguchi, N., Ota, Y., Kikuchi, N., Nobuta, K., Aoki, T., and Yokota, T.: CO₂ and CH₄ column amounts retrieval from GOSAT TANSO-FTS SWIR, Algorithm theoretical basis document, NIES GOSAT Project, 2010.

Yoshida, Y., Ota, Y., Eguchi, N., Kikuchi, N., Nobuta, K., Tran, H., Morino, I., and Yokota, T.: Retrieval algorithm for CO₂ and CH₄ column abundances from short-wavelength infrared spectral observations by the Greenhouse Gases Observing Satellite, *Atmos. Meas. Tech.*, 4, 717–734, doi:10.5194/amt-4-717-2011, 2011.